Theoretical Computer Science 134 (1994) 253–262
Elsevier

Exponential upper and lower bounds for the order of a regular language*

Andreas Weber

Fachbereich Informatik, Johann Wolfgang Goethe-Universität, Postfach 111 932, D-60054 Frankfurt am Main, Germany

Abstract

Weber, A., Exponential upper and lower bounds for the order of a regular language, Theoretical Computer Science 134 (1994) 253–262.

The order of the language recognized by a nondeterministic finite automaton with n states is either infinite or at most $3^{2n(n+1)}-1$. For each even n there is a deterministic finite automaton having n states and recognizing a language of order $2^{(n-2)/2}-1$.

0. Introduction

The order of a language L is the minimal nonnegative d such that the equality $L^* = \bigcup_{i=0}^{d} L^i$ holds or is infinite, depending on whether or not such a d exists. Using different methods the following results on the order of a regular language were shown by K. Hashiguchi, H. Leung, I. Simon, and the author. The order of the language recognized by a deterministic finite automaton with n states is either infinite or at most $2^{4n^2}-1$ ([3], see [9, Section 3.2]). This bound deteriorates to $2^{2^{2n+2}}-1$ if the automaton in question is nondeterministic [11]. It is decidable whether this order is finite ([3, 10, 11], see [9, Section 3.2]). In fact, the problem is PSPACE-complete for nondeterministic finite automata [7, 13]. For every $n \geqslant 2$ there is a nondeterministic finite automaton having n states and recognizing a language of order $2^{n-2}-1$ [13].

Correspondence to: A. Weber, Fachbereich Informatik, J.W. Goethe-Universität, Postfach 111 932, D-60054 Frankfurt am Main, Germany. Email: weber@psc.informatik.uni-frankfurt.de.

* The results of this article was presented at the Workshop on Semigroups, Formal Languages and Combinatorics on Words, Kyoto, 1992.

SSDI 0304-3975(93)E00009-8

Sections 2 and 3 present the following results.

(1) The order of the language recognized by a nondeterministic finite automaton with n states is either infinite or at most $3^{2n(n+1)}-1$ (Section 2).

(2) For each even n, there is a deterministic finite automaton having n states and recognizing a language of order $2^{(n-2)/2}-1$ (Section 3).

In order to establish result (1) we employ distance automata. A distance automaton M is a nondeterministic finite automaton which is equipped with a nonnegative cost function on its transitions. The values of this cost function are required to be 0 or 1. The distance of an input word x recognized by M is the sum of costs along a cheapest accepting path (or successful computation) consuming x. The distance of M is the maximal distance of an input word recognized by this machine or is infinite, depending on whether or not a maximum exists.

Since the order of a regular language is a special case of distance [10], we prove instead of (1) the following result.

(3) Let M be a distance automaton having n states and a unique initial and final state q_0 such that every transition having costs 1 leads to q_0. Then the distance of M is either infinite or at most $3^{2n(n-1)}-1$ (Section 2).

In our proof of (3) we use essential ideas of [3], which were further developed in [11, 9], but a different style of presentation. In order to establish result (2) we borrow a regular language from [13]. It remains open whether the bounds in (1) and (3) can be improved to the order $2^{\Theta(n)}$ or whether the bound in (2) can be improved to the order $2^{\Theta(n^2)}$. Upper and lower bounds for the distance of other distance automata than those in (3) can be found in [2, 4, 5, 8, 12–14]. The complexity of the problem to decide whether the language recognized by a given deterministic finite automaton has finite order remains open.

1. Preliminaries

We use the following notations. $\mathbb{N}$ and $\mathscr{Z}$ denote the sets of all natural numbers and of all integers, respectively. $[m]$ denotes the set $\{1,\ldots,m\}$. Throughout we assume that the model of computation for all our deterministic algorithms is the deterministic random access machine (RAM) without multiplications and divisions using the uniform cost criterion, while it is the (not necessarily always halting) nondeterministic Turing machine for all the nondeterministic algorithms (see, e.g., [1]).

A *finite distance automaton* is a 5-tuple $M=(Q,\Sigma,\gamma,Q_\mathrm{I},Q_\mathrm{F})$ where Q and Σ denote nonempty, finite sets of states and input symbols (or letters), respectively, Q_I, $Q_\mathrm{F}\subseteq Q$ denote sets of initial and final (or accepting) states, respectively, and γ is a total function $\gamma: Q\times\Sigma\times Q\to\{0,1,\infty\}$. Σ is called the input alphabet and γ is called the cost function. Each element (p,a,q) of $\gamma^{-1}(\{0,1\})$ denotes a *transition* having *costs* $\gamma(p,a,q)$. In general, M will be *nondeterministic*. Since we only deal with distance automata of the above type, the adjective "finite" is omitted from now on. M is called *deterministic* if it has exactly one initial state and if, for each $(p,a)\in Q\times\Sigma$, there is exactly one transition of the form (p, a, q). If every transition has costs 0, then M is a *finite automaton*, where we again omit the adjective "finite". The latter definition is, of course, isomorphic to the usual one of a nondeterministic finite automaton, and our deterministic automaton is the same as the usual deterministic finite automaton.

The mode of operation of M is described by paths. A *path* π (of length m) is a word $(q_1, x_1) \ldots (q_m, x_m) q_{m+1} \in (Q \times \Sigma)^m Q$ such that $(q_1, x_1, q_2), \ldots, (q_m, x_m, q_{m+1})$ are transitions. π is said to lead from q_1 to q_{m+1} and to consume $x_1 \ldots x_m \in \Sigma^*$. It is called *accepting* if q_1 is an initial and q_{m+1} is a final state. $\gamma(\pi) := \Sigma_{i=1}^{m} \gamma(q_i, x_i, q_{i+1})$ denotes the *costs* of π. By definition, $\gamma(\pi)$ is at most m. If M is an automaton, then $\gamma(\pi)$ is 0. For each $(p, x, q) \in Q \times \Sigma^* \times Q$, $d_M(p, x, q)$ is defined as the minimal costs of a path in M consuming x and leading from p to q or is infinite, depending on whether or not such a path exists. In particular, $d_M(p, \varepsilon, q)$ is 0 if $p = q$ and infinite if $p \neq q$ and, for all $a \in \Sigma$, $d_M(p, a, q) = \gamma(p, a, q)$. The *transition relation* of M, denoted by δ_M, is the set of all $(p, x, q) \in Q \times \Sigma^* \times Q$ such that $d_M(p, x, q)$ is finite. Let $\pi_1 = \pi'_1 q_1$ and $\pi_2 = \pi'_2 q_2$ be paths in M leading from p_1 to q_1 and from p_2 and q_2, respectively. If q_1 and p_2 coincide, then we define the path $\pi_1 \circ \pi_2 := \pi'_1 \pi'_2 q_2$. Note that the operation $\circ$ on paths is associative. The *size* of M, denoted by $\| M \|$, is defined as the sum of $\# Q$, $\# \Sigma$, and the number of all transitions of M.

The *language recognized* by M, denoted by $L(M)$, is the set of words consumed by the accepting paths in M. The *distance* of $x \in \Sigma^*$ in M (short form: $d_M(x)$) is defined as the minimal costs of an accepting path in M consuming x or is infinite, depending on whether or not such a path exists, i.e.,

$$d_M(x) = \min(\{\infty\} \cup \{d_M(p, x, q) \mid p \in Q_\mathrm{I}, q \in Q_\mathrm{F}\}).$$

Note that $d_M(x)$ is finite if and only if x belongs to $L(M)$. The *distance* of M (short form: $d(M)$) is the supremum of the set $\{0\} \cup \{d_M(x) \mid x \in L(M)\}$.

Let $x = x_1 \ldots x_m \in \Sigma^*$ $(x_1, \ldots, x_m \in \Sigma)$, and let $j \in \{0, \ldots, m\}$. We define

$$\mathrm{att}(x, j) := \{q \in Q \mid \exists q_\mathrm{I} \in Q_\mathrm{I}\colon (q_\mathrm{I}, x_1 \ldots x_j, q) \in \delta_M\},$$

$$\mathrm{der}(x, j) := \{q \in Q \mid \exists q_\mathrm{F} \in Q_\mathrm{F}\colon (q, x_{j+1} \ldots x_m, q_\mathrm{F}) \in \delta_M\},$$

and

$$\mathrm{set}(x, j) := \mathrm{att}(x, j) \cap \mathrm{der}(x, j);$$

$\mathrm{att}(x, j)$ and $\mathrm{der}(x, j)$ denote the set of states attainable from Q_I with $x_1 \ldots x_j$ and the set of states derivable to Q_F with $x_{j+1} \ldots x_m$, respectively (see, e.g., [9, Section 3.2] or [15, Section 2]).

The *graph of accepting paths* in M consuming x (short form: $G_M(x)$) is the directed graph (V, E) where

$$V := \{(q, j) \in Q \times \{0, \ldots, m\} \mid q \in \mathrm{set}(x, j)\}$$

and

$$E := \{((p, j-1), (q, j)) \in V^2 \mid j \in [m], (p, x_j, q) \in \delta_M\}.$$

Thus, for each $j \in \{0, \ldots, m\}$, $\mathrm{set}(x,j)$ is the set of states which appear at column j in $G_M(x)$. Let us assume that x belongs to $L(M)$ and that any edge $((p,j-1),(q,j))$ of $G_M(x)$ has costs $\gamma(p,x_j,q)$. Then the minimal sum of the costs along a path in $G_M(x)$ leading from $Q_\mathrm{I} \times \{0\}$ to $Q_\mathrm{F} \times \{m\}$ coincides with the distance of x in M. Each vertex of $G_M(x)$ is situated on such a path.

Using the above notations we define the criterion (ID) (see [3, 11; 9, Lemma 3.10]), where "ID" stands for "infinite distance".

(ID): There is a word $x = x_1 \ldots x_m \in L(M)$ $(x_1, \ldots, x_m \in \Sigma)$ and there are $0 \leqslant j_1 < j_2 \leqslant m$ such that $(\mathrm{set}(x,j_1), \mathrm{att}(x,j_1))$ and $(\mathrm{set}(x,j_2), \mathrm{att}(x,j_2))$ coincide and, for all $p, q \in \mathrm{set}(x,j_1)$, $d_M(p, x_{j_1+1} \ldots x_{j_2}, q)$ is greater than 0.

The *order* of a subset L of Σ^* is the minimal nonnegative d such that the equality $L^* = \bigcup_{i=0}^{d} L^i$ holds or is infinite, depending on whether or not such a d exists. The following proposition shows that the order of a regular language is a special case of distance.

Proposition 1.1 (Simon [10, Section 7]). *Let $M = (Q, \Sigma, \gamma, Q_\mathrm{I}, Q_\mathrm{F})$ be an automaton recognizing the language L. We associate with M the distance automaton $M' = (Q', \Sigma, \gamma', \{q_0\}, \{q_0\})$ where $Q' := Q \dot{\cup} \{q_0\}$ and, for all $(p', a, q') \in Q' \times \Sigma \times Q'$,*

$$\gamma'(p', a, q') := \begin{cases} 0 & \text{if } p' = q_0 \text{ and } \exists p \in Q_\mathrm{I}\colon (p, a, q') \in \delta_M \text{ or if } (p', a, q') \in \delta_M, \\ 1 & \text{if } q' = q_0 \text{ and } \exists q \in Q_\mathrm{F}\colon (p', a, q) \in \delta_M \\ & \quad \text{or if } p' = q' = q_0 \text{ and } \exists p \in Q_\mathrm{I} \exists q \in Q_\mathrm{F}\colon (p, a, q) \in \delta_M, \\ \infty & \text{otherwise.} \end{cases}$$

M' recognizes L^* and its distance coincides with the order of L. M' has size at most $4\|M\|$, and it can be constructed in deterministic time linear in the size of M.

Proof. The second statement of the proposition is obvious. For the proof of the first statement it is sufficient to note that, for each $x \in L(M')$, x belongs to $L^{d_{M'}(x)}$ and that, for each $x \in L^i$, $d_{M'}(x)$ is at most i $(i \in \mathbb{N} \cup \{0\})$. $\square$

2. Upper bound

In this section we prove the following theorem.

Theorem 2.1. *Let M be a distance automaton having n states and a unique initial and final state q_0 such that every transition having costs 1 leads to q_0. Then the assertions* (i)–(iii) *are equivalent.*

(i) *M has distance at most $3^{2n(n-1)}-1$.*
(ii) *M has finite distance.*
(iii) *M does not comply with (ID).*

Replacing in (ID) any occurrence of $(\text{set}(x,j),\text{att}(x,j))$ by $(\text{att}(x,j),\text{der}(x,j))$, Theorem 2.1 was implicitly proved by Hashiguchi with an upper bound of order $2^{\Theta(n^2)}$ in assertion (i) and under the assumption that M restricted to those transitions not leading to q_0 is deterministic ([3], see [9, Section 3.2]). It is not clear whether the latter assumption can be dropped in this proof. Using the same version of (ID) as Hashiguchi and some of his ideas, Simon explicitly established Theorem 2.1 with the upper bound of $2^{2^{2n}}-1$ in assertion (i) [11]. In our proof we use all essential ideas of [3], as we understand them, but with a different style of presentation.

Theorem 2.1 and Proposition 1.1 directly imply the following theorem.

Theorem 2.2. *Let M be an automaton with n states. If the order of the language recognized by M is finite, then it is at most $3^{2n(n+1)}-1$.*

Recently, the author constructed, for each $n\geqslant 2$, an automaton having n states and recognizing a language of order $2^{n-2}-1$ [13]. In Section 3, we present, for each even n, a deterministic automaton having n states and recognizing a language of order $2^{(n-2)/2}-1$. It remains open whether the bounds of Theorems 2.1 and 2.2 can be improved to the order $2^{\Theta(n)}$.

It is easy to obtain a nondeterministic polynomial-space algorithm testing whether a given distance automaton complies with (ID). Therefore, using Proposition 1.1 and Theorem 2.1, it can be decided in polynomial space whether the language recognized by a given automaton has infinite order. The same result was first proved by Leung by means of an "algebraic" algorithm [7].

In the remainder of this section we prove Theorem 2.1.

Proof of Theorem 2.1. Let $M=(Q,\Sigma,\gamma,\{q_0\},\{q_0\})$ be a distance automaton with n states such that every transition having costs 1 leads to q_0. Consider the assertions (i)–(iii) of the theorem. Clearly, (i) implies (ii). It remains to be shown that (ii) implies (iii) and that (iii) implies (i).

(ii)$\Rightarrow$(iii): Assume that M complies with (ID). Let $u,w\in\Sigma^*$ and $v\in\Sigma^+$ such that $x:=uvw$ belongs to $L(M)$, $(\text{set}(x,|u|),\text{att}(x,|u|))$ and $(\text{set}(x,|uv|),\text{att}(x,|uv|))$ coincide, and, for all $p,q\in A:=\text{set}(x,|u|)$, $d_M(p,v,q)$ is greater than 0.

Let $t\in\mathbb{N}$, and let π be any accepting path in M consuming uv^tw. Consider the uniquely determined paths $\pi_0,\ldots,\pi_{t+1}$ and states $q_0,\ldots,q_{t+2}$ in M such that $\pi=\pi_0\circ\cdots\circ\pi_{t+1}$, π_0 consumes u, $\pi_1,\ldots,\pi_t$ all consume v, π_{t+1} consumes w, and, for all $i\in\{0,\ldots,t+1\}$, π_i leads from q_i to q_{i+1}. We are going to prove by induction on t.

Claim 1. *$q_1, \ldots, q_{t+1}$ all belong to A.*

From Claim 1 it follows that $\gamma(\pi) \geqslant \Sigma_{i=1}^{t} \gamma(\pi_i) \geqslant \Sigma_{i=1}^{t} d_M(q_i, v, q_{i+1}) \geqslant t$. Since $uv^t w$ belongs to $L(M)$ and π and t were arbitrary, this implies that M has infinite distance, as desired. It remains to prove Claim 1.

Proof of Claim 1. If $t=1$, then $q_1, q_2 \in A$. Thus, let us assume that $t>1$. Since $q_2 \in \mathrm{att}(x, |u|)$, there is a path π'_0 in M consuming u and leading from Q_{I} to q_2. The induction hypothesis applied to $\pi' := \pi'_0 \circ \pi_2 \circ \cdots \circ \pi_{t+1}$ yields that $q_2, \ldots, q_{t+1}$ all belong to A. Since $q_2 \in A \subseteq \mathrm{der}(x, |uv|)$, we obtain that $q_1 \in \mathrm{att}(x, |u|) \cap \mathrm{der}(x, |u|) = A$. $\square$

Proof of Theorem 2.1 (*continued*). (iii)$\Rightarrow$(i): Assume that M does not comply with (ID). We define the function $\varphi : 2^Q \times 2^Q \to \mathscr{Z}$ by setting $\varphi(A, B) := \#A + \#B - 2$ for $A, B \in 2^Q$. We further define a complete order $\leqslant$ on $2^Q \times 2^Q$ by saying that $(A_1, B_1) \leqslant (A_2, B_2)$ if $\varphi(A_1, B_1) \leqslant \varphi(A_2, B_2)$, for $A_1, A_2, B_1, B_2 \in 2^Q$.

Let $x = x_1 \ldots x_m \in L(M)$ $(x_1, \ldots, x_m \in \Sigma)$. We define

$$\mathrm{pairs}(x) := \{(\mathrm{set}(x, j), \mathrm{att}(x, j)) \mid j \in \{0, \ldots, m\}\},$$

$$\kappa(x) := \max\{\varphi(A, B) \mid (A, B) \in \mathrm{pairs}(x)\},$$

and

$$\lambda(x) := \#\{(A, B) \in \mathrm{pairs}(x) \mid \varphi(A, B) = \kappa(x)\}.$$

Since, for all $(A, B) \in \mathrm{pairs}(x)$, $\emptyset \neq A \subseteq B \subseteq Q$, we observe that $\kappa(x) \in \{0, \ldots, 2n-2\}$ and $\lambda(x) \in \{1, \ldots, 3^n - 2^n\}$. For each $(\kappa, \lambda) \in \{0, \ldots, 2n-2\} \times \{0, \ldots, 3^n - 2^n\}$, we define $f(\kappa, \lambda)$ to be the maximal distance of a word $x \in L(M)$ such that either $\kappa(x) < \kappa$ or $\kappa(x) = \kappa$ and $\lambda(x) \leqslant \lambda$, or to be 0 if no such word exists, or to be infinite if no maximum exists. We are going to show Claims 2 and 3.

Claim 2. $f(0, 3^n - 2^n) = 0$.

Claim 3. *For all $(\kappa, \lambda) \in [2n-2] \times [3^n - 2^n]$, $f(\kappa, \lambda)$ is at most $1 + f(\kappa - 1, 3^n - 2^n) + f(\kappa, \lambda - 1)$.*

Let $\kappa \in [2n-2]$. From Claim 3 and the fact that $f(\kappa, 0) \leqslant f(\kappa - 1, 3^n - 2^n)$ it follows that $f(\kappa, 3^n - 2^n)$ is at most $3^n - 2^n + (3^n - 2^n + 1) \cdot f(\kappa - 1, 3^n - 2^n)$. Using this and Claim 2 it is easy to show by induction on κ that, for all $\kappa \in \{0, \ldots, 2n-2\}$, $f(\kappa, 3^n - 2^n)$ is at most $(3^n - 2^n + 1)^\kappa - 1$. Hence, the distance of M is at most $3^{2n(n-1)} - 1$.

It remains to prove Claims 2 and 3.

Proof of Claim 2. Let $x \in L(M)$ such that $\kappa(x) = 0$. Then, $\mathrm{set}(x, 0) = \mathrm{att}(x, 0) = \mathrm{set}(x, |x|) = \{q_0\}$ (as for any $x \in L(M)$) and $\mathrm{att}(x, |x|) = \{q_0\}$ (since $\kappa(x) = 0$). Since

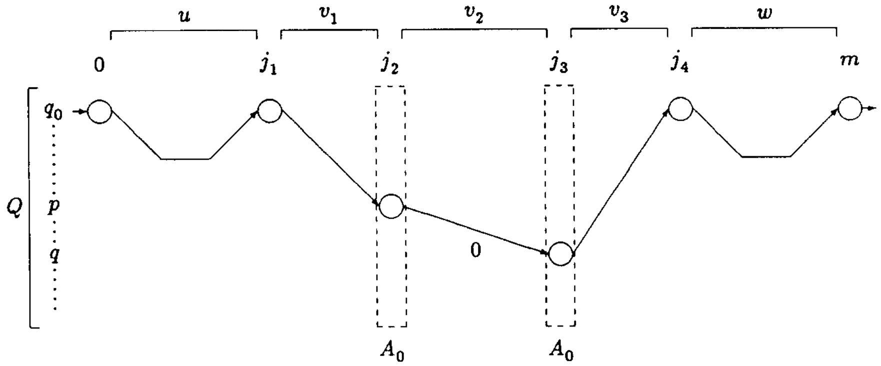

Fig. 1. Extract from $G_M(x)$.

M does not comply with (ID), this implies that $d_M(x)=d_M(q_0,x,q_0)=0$. Thus, Claim 2 is valid. □

Proof of Claim 3. Let $(\kappa,\lambda)\in[2n-2]\times[3^n-2^n]$, and let $x\in L(M)$ such that $\kappa(x)=\kappa$ and $\lambda(x)=\lambda$. We are going to show that the distance of x in M is at most $1+f(\kappa-1,3^n-2^n)+f(\kappa,\lambda-1)$. From this it follows that Claim 3 is correct.

Let $x=x_1\ldots x_m$ $(x_1,\ldots,x_m\in\Sigma)$. Note that $\varphi(\mathrm{set}(x,0),\mathrm{att}(x,0))=0<\kappa(x)$. Let j_2 be the minimal $j\in\{1,\ldots,m\}$ such that $\varphi(\mathrm{set}(x,j),\mathrm{att}(x,j))$ equals $\kappa(x)$. Let j_3 be the maximal $j\in\{j_2,\ldots,m\}$ such that $(\mathrm{set}(x,j),\mathrm{att}(x,j))$ and $(\mathrm{set}(x,j_2),\mathrm{att}(x,j_2))$ coincide. Since M does not comply with (ID), there are states p, $q\in A_0:=\mathrm{set}(x,j_2)$ such that $d_M(p,x_{j_2+1}\ldots x_{j_3},q)=0$. By definition of A_0 we can find a $j_1\in\{0,\ldots,j_2\}$ and a $j_4\in\{j_3,\ldots,m\}$ such that $u:=x_1\ldots x_{j_1}\in L(M)$,

$$d_M(q_0,x_{j_1+1}\ldots x_{j_2},p)+d_M(q,x_{j_3+1}\ldots x_{j_4},q_0)\leqslant 1,$$

$w:=x_{j_4+1}\ldots x_m\in L(M)$, and $A_0\neq\{q_0\}$ or $j_1<j_2$. We further set $v_1:=x_{j_1+1}\ldots x_{j_2}$, $v_2:=x_{j_2+1}\ldots x_{j_3}$, and $v_3:=x_{j_3+1}\ldots x_{j_4}$ (see Fig. 1).

If $j_1=j_2$ and $A_0\neq\{q_0\}$, then A_0 contains q_0 and at least one additional state and $\varphi(\mathrm{set}(u,|u|),\mathrm{att}(u,|u|))<\varphi(A_0,\mathrm{att}(x,j_2))=\kappa(x)$. Moreover, $\varphi(\mathrm{set}(w,0),\mathrm{att}(w,0))=0<\kappa(x)$. Now, the following facts are straightforward.

Fact 1. $\kappa(u)<\kappa(x)=\kappa$.

Fact 2. $\kappa(w)\leqslant\kappa(x)=\kappa$.

Fact 3. *Any* $(A,B)\in\mathrm{pairs}(w)$ *with* $\varphi(A,B)=\kappa(x)=\kappa$ *belongs to* $\mathrm{pairs}(x)$ *and is distinct from* $(\mathrm{set}(x,j_2),\mathrm{att}(x,j_2))$.

From Fact 1 it follows that u has distance at most $f(\kappa-1, 3^n-2^n)$ in M. Facts 2 and 3 imply that either $\kappa(w)<\kappa$ or $\kappa(w)=\kappa$ and $\lambda(w) \leqslant \lambda(x)-1<\lambda$. From this it follows that w has distance at most $f(\kappa, \lambda-1)$ in M. Therefore, we can estimate

$$\begin{aligned} d_M(x) &\leqslant d_M(u)+d_M(q_0, v_1, p)+d_M(p, v_2, q)+d_M(q, v_3, q_0)+d_M(w) \\ &\leqslant 1+f(\kappa-1, 3^n-2^n)+f(\kappa, \lambda-1). \end{aligned}$$

This completes the proof of Claim 3 and of Theorem 2.1. □

3. Lower bound

In this section we prove the following theorem.

Theorem 3.1. *For each $n \geqslant 2$ there is a deterministic automaton M_n having $2n$ states, $2n-2$ input symbols, and $\Theta(n^2)$ transitions and recognizing a language of order $2^{n-1}-1$.*

Using a simple binary coding, we may diminish the number of input symbols of M_n (in Theorem 3.1) to 2, while the order of the language recognized remains unchanged and the number of states is increased to $O(n \log_2 n)$. In order to prove Theorem 3.1, we just borrow from [13, Theorem 2.2] a regular language of order $2^{n-1}-1$ and we verify that the minimal deterministic automaton recognizing that language has the desired properties. It remains open whether there are regular languages as in Theorem 3.1 having order $2^{\Theta(n^2)}$.

In the remainder of this section we prove Theorem 3.1.

Proof of Theorem 3.1. Let $n \geqslant 2$, and let $a_1, b_1, \ldots, a_{n-1}, b_{n-1}$ be pairwise distinct letters. We define $\Sigma_j := \{a_1, b_1, \ldots, a_j, b_j\}$ $(j=0, \ldots, n-1)$ and $L := \{a_1\}(\Sigma_1)^*\{b_1\} \cup \cdots \cup \{a_{n-1}\}(\Sigma_{n-1})^*\{b_{n-1}\}$. It is shown in the proof of [13, Theorem 2.2] that L has order $2^{n-1}-1$. On the other hand, this language is recognized by the deterministic automaton $M_n = (Q, \Sigma, \gamma, Q_I, Q_F)$ where

$$Q := \{q_1, \ldots, q_{2n}\}, \quad \Sigma := \Sigma_{n-1}, \quad Q_I := \{q_2\}, \quad Q_F := \{q_3, q_5, \ldots, q_{2n-1}\},$$

$$\gamma(q_{i_1}, a_j, q_{i_2}) := \begin{cases} 0 & \text{if } i_1=2 \text{ and } i_2=2j+2 \\ & \quad \text{or if } i_1>2j \text{ and } i_2 \in \{i_1, i_1+1\} \text{ is even} \\ & \quad \text{or if } i_1 \in \{1, \ldots, 2j\} \setminus \{2\} \text{ and } i_2=1, \\ \infty & \text{otherwise,} \end{cases}$$

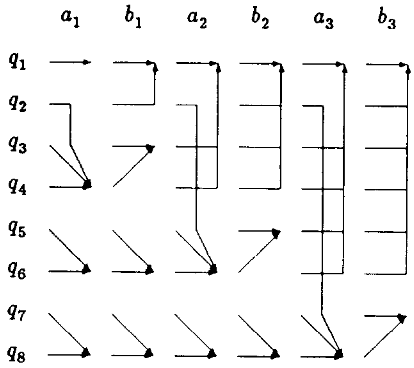

Fig. 2. Definition of γ in M_n ($n=4$).

and

$$\gamma(q_{i_1}, b_j, q_{i_2}) := \begin{cases} 0 & \text{if } i_1 \in \{2j+1, 2j+2\} \text{ and } i_2 = 2j+1 \\ & \quad \text{or if } i_1 > 2j+2 \text{ and } i_2 \in \{i_1, i_1+1\} \text{ is even} \\ & \quad \text{or if } i_1 \in \{1, \ldots, 2j\} \text{ and } i_2 = 1, \\ \infty & \text{otherwise} \end{cases}$$

($1 \leqslant i_1, i_2 \leqslant 2n$, $1 \leqslant j \leqslant n-1$, see Fig. 2).

M_n has $\# Q \cdot \# \Sigma = 4n^2 - 4n$ transitions. Using standard formal language theory (see [6]) it is straightforward to check that M_n is minimal. $\square$

References

[1] A. Aho, J. Hopcroft and J. Ullman, *The Design and Analysis of Computer Algorithms* (Addison-Wesley, Reading, MA, 1974).

[2] P. Gohon, Automates de coût borné sur un alphabet à une lettre, *RAIRO Inform. Théor.* **19** (1985) 351–357.

[3] K. Hashiguchi, A decision procedure for the order of regular events, *Theoret. Comput. Sci.* **8** (1979) 69–72.

[4] K. Hashiguchi, Limitedness theorem on finite automata with distance functions, *J. Comput. System Sci.* **24** (1982) 233–244.

[5] K. Hashiguchi, Improved limitedness theorems on finite automata with distance functions, *Theoret. Comput. Sci.* **72** (1990) 27–38.

[6] J. Hopcroft and J. Ullman, *Introduction to Automata Theory, Languages, and Computation* (Addison-Wesley, Reading, MA, 1979).

[7] H. Leung, An algebraic method for solving decision problems in finite automata theory, Ph.D. Thesis, Pennsylvania State University, 1987.

[8] H. Leung, On finite automata with limited nondeterminism, submitted; see also: *Proc. MFCS 1992*, Lecture Notes in Computer Science Vol. 629 (Springer, Berlin, 1992) 355–363.

[9] A. Salomaa, *Jewels of Formal Language Theory* (Pitman, London, 1981).

[10] I. Simon, Limited subsets of a free monoid, in: *Proc. 19th IEEE FOCS* (1978) 143–150.
[11] I. Simon, On Brzozowski's problem: $(1 \cup A)^m = A^*$, in: M. Fontet and I. Guessarian, eds., *Seminaire d'Informatique Théorique 1979-1980*, LITP, Université Paris VI/VII (1980), pp. 67–72.
[12] I. Simon, On semigroups of matrices over the tropical semiring, Tech. Report RT-MAC-8907, IME, Universidade de São Paulo, 1989.
[13] A. Weber, Distance automata having large finite distance of finite ambiguity, *Math. Systems Theory* **26** (1993) 169–185.
[14] A. Weber, Finite-valued distance automata, *Theoret. Comput. Sci.* **134** (1994) 225–251, this volume.
[15] A. Weber and H. Seidl, On finitely generated monoids of matrices with entries in $\mathbb{N}_0$, *RAIRO Inform. Théor. Appl.* **25** (1991) 19–38.